The semantics of case[*]

ANNABEL CORMACK

1 Introduction

As it is currently understood within P&P theory, the Case module appears to be a purely syntactic condition, contributing to regulating the syntactic distribution of argument noun phrases, and, in some versions of the theory, argument clauses. In most, or perhaps all, other theories of grammar, there is no equivalent. This paper will argue for the position that abstract Case is not simply a syntactic matter, but has reflexes in both lexical and compositional semantics.

Given some particular principles concerning the projection and discharge of θ-roles, in the spirit of Higginbotham (1985 etc.) and Williams (1994, etc.), Case can be seen as a simple but powerful means of regulating the parallel workings of syntax and semantics. The investigation of the operation of Case from a perspective which views syntax as semantically transparent leads to two simple theses. First, Case is implicated in lexical semantics. The lexical reflex of Case lies in the existence of distinct [+Case] and [-Case] entries for the canonic argument heads D and C and P. The [+Case] version heads a (semantic) argument, but the [-Case] version must head a (semantic) predicate. Secondly, Case is implicated in compositional semantics. The compositional reflex of Case is instantiated in the variable semantic content of AGR, which is required to license the discharge of θ-roles. If the relevant projection of the selecting head has the [+Case] feature, AGR must contain the applicative combinator **A**; this induces function-argument application to discharge the theta role.[1] When the head projects with [-Case], the content of AGR must be the composing combinator **R**. The effect of this **R** combinator is to allow θ-roles to be saturated non-locally, mimicking Raising, without invoking either movement or chains. That is, in

[*]This is a shortened version of part of a longer paper in preparation (Cormack 1995). I would like to thank the Leverhulme Trust for a grant which allowed me to have Richard Breheny as a Research Assistant. Without his help, the papaer could not have been written; without our discussions, it would not have been such fun. Yorgos Xydopoulos has also assisted me, financed by an extension to the grant, and I am grateful for his help. I also wish to thank Irene Heim, Hans Kamp, Ruth Kempson, Rita Manzini, and Neil Smith for a number of useful questions, criticisms and suggestions. I am also indebted to Mark Steedman: the influcence of generalised and combinatory categorial grammar should be apparent. The errors are mine.

[1]For combinators, see Steedman (1989), and Steedman (1988) for **W′**, which I have re-christened **R**. For **A**, see Szabolcsi (1990).

minimalist terms (Chomsky 1995), A-movement is to be seen as an instance of Merge, rather than as an instance of Attract/Move.

These two Case effects not only give an account of the semantics of the familiar cases of A-chains, but have further interesting consequences. The θ-position trace can be given a lexical entry complete with semantics. The composing combinator turns out to be implicated in the semantics of ECM and Small Clause constructions, with the concomitant syntactic consequence that a 'Raising to Object' analysis is necessary. Some instances of Control structures can be seen to fall under the same analysis (though I shall not argue for that here).

The paper will be structured as follows. Section 2 introduces the background assumptions which drive the rest of the analysis. This section is not directly concerned with Case at all, but sets out and produces some arguments for the principles and notation which will be used later. The argument proper begins in section 3, which concerns the lexical semantics of Case as it applies to determiners. Section 4 is about the compositional reflexes of Case, motivated by Raising. Section 5 discusses Small Clauses. The next section, 6, puts the proposals in perspective. Section 6.2 completes the inventory of AGR by adding AGRs. It is proposed in section 6.3 that it is redundant to check Case on D. Section 6.4 gives a Minimalist account of the categorial features involved in the two kinds of Merge that have been argued for. Section 7 offers a conclusion.

2 Predicates, operators, and binders

2.0 This section is not directly concerned with Case at all, but sets out and briefly argues for the principles and notation which will be used later. It is concerned with predicates, functional heads and quantified noun phrases.

The initial section is simply about notation, and the interpretation of XP. Section 2.2 discusses how functional heads are projected, considering both one- and two-place operators, such as *not* and *because* respectively. Section 2.3 characterises a determiner as a two-place binding operator, giving a new form to the DP hypothesis. The final subsection discusses the s-selection of arguments with nil semantic content, for example in the external argument position of an unaccusative verb. This innovation turns out to have considerable consequences.

With respect to syntax, I will be as conservative as is consistent with my claims. I believe that these are compatible with the Minimalist program (Chomsky 1991, 1993, 1994, 1995), but I have not systematically cast them in that framework. So far as semantics is concerned, I will use the minimum apparatus which will enable me to demonstrate the claims I make. It should be borne in mind, then, that there is a good

deal of simplification.[2] As a basis for the semantic types, I use just <*e*>, for entity, and <*t*>, for proposition (truth-value bearer). The type system registers the resulting type for the mother, as well as the types for the daughters.

2.1 Predicates and non-binding functional heads

In order to talk about subject θ-roles in a perspicuous fashion, I am reverting in this paper to the older notation (Chomsky 1981, for instance) under which the projection XP of a lexical (relational) head X is a predicate. By a predicate, I mean a projection which could semantically function as the predicate of an external subject, whether or not it actually does so function. A predicate, then, is a projection with an undischarged external θ-role.

The specifier position within XP can usefully be utilised to show the presence of such a role, as in (1)[3]:

(1) $[_{NP}\ [_{DP}\ \theta]\ [_{N}\ \textit{dog}]]$ type <*e,t*>
$[_{VP}\ [_{DP}\ \theta]\ [_{V'}\ [_{V}\ \textit{eat}]\ ...]$ type <*e,t*>

The specifier here, with its DP and θ, is a convenient fiction, representing the fact that a projection with lexical content of head (and complement) still projects an unsaturated θ-role.[4] This θ-role is the external role: by this is meant simply that it is the final θ-role projected. The notation embodies the decision to reject the VP-internal subject hypothesis as it is currently understood, and XP-internal subjects for other relational heads such as N. This position is close to that of Williams ((1994) and earlier papers).

The linear order in the trees above is to be taken as conventional, rather than real. If the surface variation of the order of constituents in a language is to depend solely on properties of functional elements (Chomsky (1991),[5] Borer (1984)), then UG must

---

[2] In particular, I abstract away from matters pertaining to intensionality and scope in this paper.

[3] In section 6.4, I discuss a notation based on Chomsky (1994), which uses selection features rather than a DP in the specifier position to register the external categorial selection. The external type selection is carried in the type for the projection.

[4] If the projection of the external θ-role is not made explicit, it can be seen that what we propose is related to Hellan's (1991) proposals for a two-level X′ system.

[5] Chomsky (1991) allows also for generalisations over the lexicon: it seems preferable to do without such a repository of grammatical information if possible.

either offer a fixed order of relational head to complement, as argued by Kayne (1994) and modified by Chomsky (1994), or leave the head and its complement unordered unless a functional head intervenes. In Chomsky (1981 p 94), it was argued that the base should have no order stipulated; this line is followed here. (The reader is to supply the variant trees, as necessary).

2.2 Functional projections

In Cormack and Breheny (1994), we offer an alternative to the standard notion of an adjunct. We argue that functional heads (including minor categories) project as features, with the maximal projection of the functional head bearing the category of the last operand.[6] They form double-headed projections. I offer a Minimalist revision of this notation in section 6.4 below.

Consider a one-place non-binding operator like *not*, of functional category F,[7] taking an AP as its operand. We would have *not cold* projecting as an AP with the feature [FP] projected from the head F, as in (2):

(2) $[_{AP[FP]}$ *not* $[_{AP}$ θ *cold*]] type <*e,t*>

The type for *cold* is <*e,t*>,[8] and if the type for the operator *not* is <<*e,t*>,<*e,t*>>, function argument application returns a type for the whole constituent which is still <*e,t*>. It is typical of non-binding operators that the input and output types are the same.[9] This whole constituent functions as a predicate, since the θ-role from *cold* is still unbound .

The standard functional heads like Infl will be one-place operators, so that since Infl selects VP for its operand, instead of IP we would have VP[IP].

---

[6]I restrict the term 'argument' to categories that saturate θ-roles projected by relational heads. Functional heads will not be said to project θ-roles, but rather to demand operands. The syntactic transparency of the head F with respect to its operand is characterised by Jacobson (1990) as Lexical Inheritance.

[7]The category F represents any functional head where the particular category is unimportant here.

[8]I ignore the fact that the adjective is in fact unaccusative (Cormack and Breheny (1994)).

[9]I am assuming here that they are also polymorphic. That is, if the canonic type for say *not* is <*t,t*>, other types such as the <<*e,t*>,<*e,t*>> required here can be systematically derived.

We also argue that two-place operators project as features, as in (3). *Because* takes two operands, which we will suppose to be a CP and an IP.[10]

(3) $[_{IP[FP]}\ [_{CP[F']}\ [_{F}$ *because*$]\ [_{CP}$ *it is raining*$]]\ [_{IP}$ *Mary is sad*$]\ ]$
Because is a two-place operator with canonic category <*t*,<*t*,*t*>>.

This analysis allows adjunct structures to conform to the principle that all structure is head-mediated. That is, the 'adjunct' constituent CP[F′] is a sister of the lower IP by virtue of the selection properties of the head F. Notice that I am claiming that the maximal projection of F is IP[FP], which does NOT dominate JUST the 'adjunct' [*because it is raining*] - this is only an intermediate level projection of F. The maximal projection of F must include the second operand, [*Mary is sad*].

Note that there are two kinds of functional projections hypothesised: one-place and two-place operator heads, and the F′ projection of a two-place operator F. A projection containing just FP is lexical. An unsaturated projection of a functional head - i.e. F or F′- is a functional category.

2.3 Binding operators: a re-interpretation of the DP hypothesis

All θ-roles must be discharged (in the sense of Higginbotham (1985)). One way of discharging a role is to bind it semantically by means of a determiner. A determiner is a functional head, but unlike those considered in the last section, it is a binder.

Consider a quantifying determiner like *every*. Natural language determiners are semantically two-place binding operators: in *every dog barks*, *every* binds the external role in both *dog* and *bark* (Barwise and Cooper 1981).[11] The unmarked assumption is that syntax follows semantics, so the determiner should have two predicate operands - the NP, and some other predicate XP. The relevant Meaning Postulates need to be able to refer to both operands. The fact that the determiner c-selects for

[10]At least as late as the end of the 18th century, the clause following *because* was introduced by the complementiser *that*. The second operand must be IP, however, since the whole IP[FP] can be embedded under *that*. The tree here is for expository purposes only: the structure is in fact more complex than this (see Cormack and Breheny (1994) for some discussion).

[11]If you think of *dog* and *bark* as denoting sets (the set of dogs, and of barking things, respectively), then *every* asserts that the *dog* set is a subset of the *bark* set.

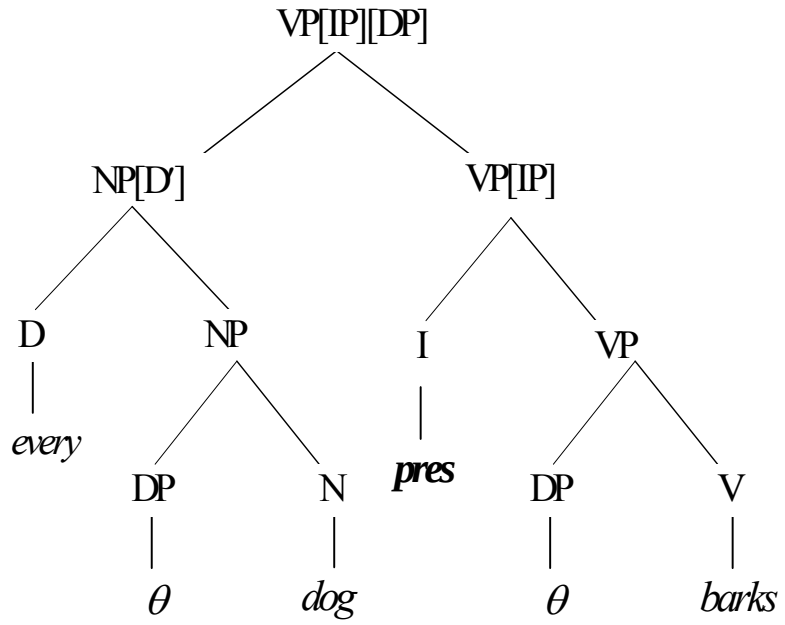
two operands allows for syntactic statements specifying Downward Entailing environments for Negative Polarity Items in English (Ladusaw 1980).[12]

So, for example, the tree for the clause in (4),

(4) Every dog barks

will be as shown in (5), where the second operand of the determiner *every* is a VP[IP].

(5)



every is a two-place binding operator of type <<e,t>,<<e,t>,t>

As is clear from the tree, the binding of the two operands by D is done in two stages: first *every* binds *dog*, to form the projection NP[D′], which is a generalised quantifier. This is the constituent usually referred to as a noun phrase. This category in turn binds the θ-role from *barks*, to give the maximal projection of D, VP[IP][DP].[13] As with the

[12]For example, the determiner *every* licenses NPIs in its first operand only, whereas the determiner *no* licenses NPIs in both operands. Since c-command as well as semantic scope seem to be involved in the licensing of NPIs in English, it is probably necessary to give the account in syntactic terms.

[13]In section 6.4, a more articulated category is proposed, under which the fact that the external role of the NP is not fully discharged until the DP level is formalised by the occurrence of N at the DP

projection of two-place operators like *because*, above, the claim is that the maximal projection of D does NOT include JUST the noun phrase [*every dog*], but must include the second operand, [*barks*], as well.

The format above allows a straightforward interpretation of a tree containing a quantifying determiner, without using the apparatus of LF, QR, OR, variables, or coindexing.[14] Furthermore, the linear order of the determiner with respect to its two operands can be given as a property of the functional category D: in English, D selects both its operands to the right, so that subjects are clause-initial, unless there is movement.

Our analysis of the relation between subject and predicate is closer to the spirit of Williams' notion of predication (Williams 1994 and earlier work), or to Higginbotham's θ-discharge, than to the current P&P analysis. It can be seen as part of a reworking of the murky notion of specifier. The subject noun-phrase, NP[D′] is in a position which is more like that of an adjunct, as argued for by Hellan (1991), Kayne (1994), and Manzini (1995). Chomsky (1994) complains that Kayne loses the distinction between adjuncts and specifiers. However, a generalised quantifier like *every* differs from an ordinary two-place operator like *because* in that it binds the θ-roles of its operands NP and AP, whereas *because* does not bind any θ-roles in its operands. We thus preserve the distinction between adjuncts and specifiers, at least in the type system (for categories, see also section 6.4).

The fact that the noun phrases associated with binding determiners are akin to adjuncts, in that they form functional categories, should explain why there is no extraction from within NP of object noun phrases.

(6) * Who did John make [the claim that he admired *t*]?

It should be clear from this description of the DP that I must reject the 'subject internal to VP' analysis in the usual form. My position is in many ways close to that of Williams. Williams (1994) summarises the objections that he has made in earlier papers to the arguments for the internal subject hypothesis. I agree with his rebuttals. Like Williams, I argue that the grammar provides mechanisms for saturation of this role by an argument which is not in the actual position at which the role is projected. Cormack (1995) adds three more arguments to those of Williams.

2.4 Nil roles

---

level as well.

[14]For object noun phrases, see section 3.2. I leave open here how semantic scope is to be related to syntax.

I assume, following the arguments in Cormack and Smith (1994) and Cormack and Breheny (1994) that an unaccusative verb like *come* projects an external role, whose semantic contribution is nil. This is shown as θ_{nil}, for mnemonic convenience. I shall assume in this paper that the external argument is always DP. Thus the VP for an unaccusative verb like *come* will be as shown in (7).

(7) $[_{VP}\ [_{DP}\ \theta_{nil}]\ [_{V'}\ [_{V}\ come]\ ...]$ type <*e,t*>

An argument whose semantic content is nil is simply one to which no reference is made in the meaning postulates for the head.[15] For example, suppose that the internal language, Fodor's Language of Thought, contains a one-place predicate *Come′* meaning what *come* would mean if it were a one-place predicate. Then we can give the meaning of the English ergative *come* as

(8) $come' = \lambda x \lambda y\ Come'\ x$

The meaning postulates associated with *come′* will be based on those of C*ome ′*, a one-place predicate, but with information relating only to the bearer of the internal role. For this reason, nil roles do not have either type <*e*> or type <*t*>, but are rather type-neutral.[16]

A nil role, of course, cannot be the sole θ-role assigned to an argument. A sentence like (9) is unacceptable, where *John* has only a nil role.

(9) John came the letters.

Such a sentence would be ruled out pragmatically, since it would flout principles of Relevance to no effect (Sperber and Wilson 1986). Whether the sentence should be characterised as syntactically deviant as well is unclear: and apparently such a characterisation would have to refer to the θ-types, if not to the meaning postulates from which these are derived. I will leave this open, noting that a similar problem arises in characterising the ungrammaticality of sentences like (10), which, as argued by Giorgi (1991), arises from the inability of a locative phrase to provide the sole θ-role for an argument.

---

[15]I am using the term 'meaning postulates' to cover both constraints set on the model, and the inference rules used by humans in processing natural language - these last being formulated over structures of the language of thought.

[16]The variable *y* in (8) is not constrained to be of any particular type, since no meaning postulates refer to it.

(10) John seems in the garden

The use of semantically nil roles means that unaccusative verbs project like transitive verbs, so that even under Chomsky's Bare Phrase Structure, such a verb would be distinct from an unergative projecting a single role. The effect of the nil roles on 'movement' is discussed in section 4.1.

3 Lexical semantics: the semantics of determiners

3.0 The argument proper begins in this section, which concerns the lexical semantics of Case as it applies to determiners. It is argued in section 3.1 that there are distinct entries in the lexicon, depending on whether the D is Case-marked or not. A [+Case] determiner is a two-place binding operator; a [-Case] determiner is a one-place non-binding operator. The several varieties of determiner are examined, to see for each what lexical entry is required for the [+Case] and [-Case] versions. Not every determiner has both entries. The determiners discussed include some phonologically empty ones. In section 3.2, a polymorphic variant of the [+Case] determiner, for object positions, is introduced. Section 3.3 summarises.

3.1 Binding and non-binding determiners

Consider the contribution to meaning made by *no* in

(11) Mary considers no fool employable

(12) Mary considers John no fool

It is generally agreed that in (11), [*no fool*] is an argument, and that in (12), [*no fool*] is rather a predicate (Chomsky (1986), p 95) In a theory without Case, this distinction is a consequence of the selection by *consider* of two complements: a noun phrase and some predicate phrase, where the latter happens to be an adjective phrase in (11), but a noun phrase in (12). In P&P theory, *consider* s-selects for a single proposition, in line with the s-selection in (13).

(13) John considers [that the fish is cooked]

The differing status of the phrase [*no fool*] must be carried by the fact that it is Case-checked in (11), but not Case-checked in (12). Let us consider this in more detail, paying attention to the compositional syntax and semantics. I assume for the purpose of exposition here a standard account under which there is a Small Clause complement. In section 5, I offer an alternative account, which takes proper account of the source of the Case.

Under earlier accounts, Case was essentially a property assigned to phrases. Then either [+Case] would turn predicates into arguments, or [-Case] would turn arguments into predicates. Given the sharp distinction between the two, this idea seems improbable. If we consider a simple semantic representation of the meaning of the phrase in each instance, the improbability mounts. In (14), the representation is the standard predicate calculus form. In (15), the simpler forms with variable-free generalised quantification and polymorphic negation are used.

(14) a) + Case: $\lambda P(\neg\exists x [\text{fool}'(x) \wedge P(x))$
b) - Case: $\lambda x(\neg \text{fool}'(x))$

(15) a) + Case: $\lambda P (\neg_2 : \text{fool}' ; P)$
b) - Case: $\neg_1 \text{fool}'$

It is clear that any relation there is between these two meanings depends on the contribution of the determiner: for instance there needs to be some relation between the two-place operator $\neg_2$ in (15a) and the one-place operator $\neg_1$ in (15b). This suggests that +/- Case variation is NOT a property of the phrase, but of the determiner *no*. Under a feature-checking account, this would be more natural, since the feature would appear in the lexicon.

Then the structure for the argument phrase [*no fool*] in (11) is the same as for [*every dog*] in (5). So, the *no* is a two place binding operator. That is, first it c-selects for an NP, the restrictor predicate, here [*fool*], and constructs a generalised quantifier, and then it looks for a second predicate, here the AP [*employable*], and binds the open argument position of that. It s-selects two predicates, i.e. phrases with open argument positions; and the final result is a complete proposition. In Montagovian simplified types, the D must have type $\langle\langle e,t\rangle,\langle\langle e,t\rangle,t\rangle\rangle$.

But now consider the same phrase as it occurs in (12). What is the effect of *no* here? It first selects a predicate as before, i.e. [*fool*]; but it must simply return a new predicate meaning (roughly, 'non-fool'). It is a one-place operator; and because it returns a predicate rather than a proposition, it is not a binder. The simplest type it can have is $\langle\langle e,t\rangle,\langle e,t\rangle\rangle$. For [*no fool*] in (12), *no* is a one-place operator much

like *not* in (2). In other words the non-binding determiner is transparent to the θ-role from the NP. The structure will be as shown in (16):

(16) $[_{NP\,[DP]}\,[_{D}\,no\,][_{NP}\,[_{DP}\,\theta\,]\,[_{N}\,fool]]]$

Given the entirely different c-selection and s-selection properties of the two kinds of occurrence, the obvious move is to postulate two entries for *no* in the lexicon. The first c-selects for NP and XP, where XP is a predicate, and is a two-place binding operator. The second c-selects just for NP, and is a one-place non-binding operator. Somehow, we need to ensure that the binder occurs only when Case-licensed. For the moment, we may suppose this to be done by means of Chomsky's checking features (Chomsky 1993). The [+Case] feature of the determiner must be checked against a Case-licensing head, otherwise the derivation will crash.

There are of course other determiners, both quantifiers and articles, which are capable of heading phrases which are arguments, and also capable of heading phrases which are predicates. These are the 'weak' determiners of Milsark (1977), more usefully described as the cardinal determiners in the terminology of Barwise and Cooper (1981). For each of these, *some*, *three*, *many*, and so on, we will again need two lexical entries, with the choice depending on Case. The binder will have a meaning given by $\lambda P\lambda Q(|P' \cap Q'|=n)$, where P' and Q' are the sets corresponding to P and Q respectively, and *n* is the cardinal. The one-place operator will have a meaning given by $\lambda P\lambda X(X' \subseteq P' \wedge |X'|= n)$.[17]

There is also at least one quantifier which is not capable of heading a predicate phrase:

(17) * Mary considers her papers [most good work on dandelions]

Then there is no entry in the lexicon for *most* without Case licensing.[18] Definite determiners normally head argument phrases, and can only head predicates under certain circumstances, but there are grammatical sentences like

(18) Mary considers John [everything that a girl could wish for].

(19) Mary considers John [the best thing since sliced bread].

---

[17]For more details concerning the semantics of these and other determiners, see Cormack (1995).

[18]There are sentences like (i) *There are most people in the garden*, but this is not synonymous with *Most people are in the garden*. The *most* in (i) is a different item, possibly a superlative adjective, which occurs in (ii) *I've got (the) most apples*.

There is also arguably at least one determiner which must not bear Case. Consider

(20) Rufus was crowned king

(21) ?* Rufus is teacher

(22) * New king was crowned.

Following Cormack and Breheny (1994), the structure in (20) will include a composite predicate [*crowned king*] which consists of the asymmetric conjunction of two unaccusative predicates, [*crowned*] and [*king*]. But in English, no NP occurs without a selecting determiner, so I assume that [*king*] is an NP[DP] with an empty determiner.[19] The semantics of this determiner asserts uniqueness but not the expectation of an already salient discourse referent. The noun phrase headed by this empty determiner cannot occur in a Case-licensed position, as we see from (22).

Let us suppose then that there is a single lexical entry for each determiner, but that for some, there is both a [+Case] and a [-Case] version within that entry, whereas for others there is only one of these. In most instances, there is a systematic relation between the two meanings involved (see Cormack (1995) for discussion).

There is a clear difference between noun phrases headed by [+Case] and [-Case] D, as witness the difference in the acceptability of extraction in (23) and (24).[20] The non-binding [-Case] D is a simple one-place operator, and does not prevent extraction in (24b), but the binding D in (23b) puts the NP into an adjunct-like position (see section 2.3).

(23) a) John broke the lid of the jar
b) ?? Which jar did John break the lid of ?

(24) a) This is the lid of the jar
b) Which jar is this is the lid of?

A similar contrast is exhibited by (25a), and (25b) (modelled on Ingham (1991)):

(25) a) Which mountain was Tensing the first man to climb ?
b) * Which mountain did you meet the first man to climb ?

---

[19]See Stowell (1991) for discussion of these issues.

[20]There appear to be dialect differences concerning the acceptability of examples like (23b).

The only relevant difference between these two is that the matrix verb in the first does and in the second does not, Case-license the noun phrase it selects as complement.[21]

The Visibility Condition of Chomsky (1981) states that every non-expletive argument noun phrase must be Case marked, before it is 'visible' for θ-role assignment.[22] We can now look at this differently. The [+Case] feature is identified with D being a binding operator, which is not an invariant property of determiners. If the [+Case] feature is absent, the determiner will head a predicate phrase, and cannot discharge a θ-role.

3.2 Object noun phrases

I have mentioned above that functional categories are frequently polymorphic. I need to appeal to this notion in accounting for object phrases (see Emms (1990) for arguments to this effect in a Categorial Grammar framework). An object binds the θ-role projected by a transitive verb. The transitive verb will have type <e,<*e,t*>>; after this role is bound, the projection has only one role available for binding, so it is now of type <*e,t*>. This entails that the noun phrase binder, the NP[D′] constituent, must be of type <<e,<*e,t*>>,<*e,t*>>, rather than the type <<*e,t*>,*t*> appropriate for subjects. The determiner itself will have type <<*e,t*>,<<e,<*e,t*>>,<*e,t*>>>, to accommodate the NP operand.[23] Its semantics will be derived from the canonic meaning as shown in (26). For the general case, see Emms (1990).[24]

(26) If the meaning of a determiner D is given by Δ, then the meaning for the object noun phrase version, Δ′, is given by
$\Delta' = \lambda P \lambda R \lambda y\ (\Delta\ P)[\lambda x\ (Rx)y]$

Given that a determiner selects both operands to the right in English, we predict that the object noun phrase is followed by the verb. This is a consequence of ordering imposed by the functional head D, not by the relational head V. I assume that a

---

[21]It is assumed that *be* in (25a) is a Raising verb, selecting as if for a Small Clause.

[22]Reformulations in terms of chains are not relevant here. For *wh*-chains, the trace will be the Case-marked argument; A-chains I have argued do not exist in the normal sense.

[23]The type-shifting determiner for proper nouns must also have polymorphic variants for object positions.

[24]The semantics for object noun phrases is in a variable-free notation. Adapting to a form with variables would give an appropriate semantics for May's (1985) adjunction to VP.

relational head is incapable of imposing any order relation between itself and a selected complement. However, I do assume that all binding operations are mediated by an AGR projection, and that in English, the verb is constrained to move to a head-initial AGRo. For some empirical evidence for this, see Cormack and Breheny (1994). The analysis suggested here owes much to Larson's analyses in his (1988a) and (1988b). The difference is that where Larson has a shell verb position, we postulate an AGR. AGR is to be generated optionally, but I propose below that an AGR is required to license any θ-role discharge.

Crucially, I assume that the subject θ-role is not and cannot be projected until the object θ-role has been bound AND licensed - a process necessarily mediated by AGR. The idea that licensing should be local is natural within the minimalist framework. The idea that a θ-role cannot be projected until the previous role is bound is a consequence of treating heads as Curried semantic or syntactic functions, as is done in categorial grammar, and in most formal semantics.[25] It is implicit in the type-notation we have been using, and in section 6.4 I will give a Minimalist version for the syntactic categories. Briefly, I assume that c-selection features form a hierarchy, realised as a stack where just one such feature is visible and available at the top of the stack at any one time. If the top feature of the stack is checked and deleted, then the next one becomes visible. Returning to a notation which puts θ for an available θ-role, a transitive VP will then be projected as in (27):

(27) $[_{VP}\ \theta_1\ [AGRo\ [NP[D']\ [_{V'}\ \theta_2\ \ V]]]]$

Here, the verb initially projects an object θ-role, θ_2. This is bound by the noun-phrase NP[D′], and after the discharge has been checked by AGRo, the subject θ-role, θ_1, is projected. (Note that in order to show explicitly the θ-role associated with the object argument, θ_2, we have V′ where we usually just have V. The bar-level notation is not particularly well-suited to the theory; and in any case, the θ-categories are notional). AGRo will also check that φ-features and morphological case match on D and the V′ projection of V. The format for a verb and a complement, with a licensing AGR, can be repeated any number of times. For example, if the verb has two Case-licensed internal arguments, then for each there will be an AGRo phrase which checks Case and morphological case, and to which the verb moves, as in (28):

(28) $[_{VP}\ \theta_1\ [AGRo_2\ [np_2\ [_{V'}\ \theta_2\ \ [AGRo_3\ [\ np_3\ [_{V''}\ \theta_3\ V]]]]]]]$

[25]Bowers (1993) similarly uses the type system to determine when a role is available for discharge, and abandons the Subject Internal Hypothesis. However, his AGRo licensing is, as in the standard P&P theory, at LF, and not local.

The AGR projections, together with the principle that θ-roles are not projected until a previously projected role has been discharged and licensed, will have much the same effect as the Larson V-shell hypothesis (Larson 1988a). The proposal minimises any difference in status between internal arguments. The consequent expectation that an indirect object could trigger verbal agreement, because of its own AGR, is met, by Georgian (Anderson (1992), p 144), and Basque (Saltarelli (1988)).

I discuss the rejection of the traditional reasons for distinguishing inherent and structural Case (with respect to internal arguments) in section 5.

3.3 Summary

In this section, it has been argued that determiners may have two lexical entries, one associated with the feature [+Case], and one with [-Case]. A [+Case] determiner is a two-place binding operator, which binds the available θ-roles in its first operand, an NP, and in its second operand. Because a second operand is still to come, the noun phrase headed by a binding determiner will have the functional category NP[D′]. These determiners have type $<<e,t>,<<e,t>,t>$, or a polymorphic variant of this. Determiners which are [-Case] are one-place operators, heading a noun phrase of category NP[DP]. They are not binders. Rather, they transmit the unbound θ-role from their operand NP: they have type $<<e,t>,<e,t>>$.

The properties of the two classes of determiners were argued for primarily on semantic grounds, but there are syntactic reflexes too: *wh*-extraction differentiates between noun phrases of the two kinds.

4 Predicates and composition

4.0 This section introduces the compositional reflexes of Case, motivated by Raising. I argue that subject np-traces are unnecessary and indeed impossible, and that instead, we should allow that a head can be combined with its complement by an adapted form of composition of functions, more or less as argued for by Jacobson (1990). Because of nil roles, where Jacobson's composition is equivalent to using the combinator **B,** it is necessary instead to use **R**.[26] This combinator essentially involves the equivalent of movement into a θ-position: section 4.2 argues that such 'movement' is in any case necessary. In section 4.3, it is argued that θ-discharge is mediated by AGR. If AGR checks a [+Case] projection of a θ-assigning head, then the associated semantics is

[26]This is the **W′** of Steedman (1988), which I have re-christened**.** See references of footnote 1.

function-argument application: the content of AGR is the applicative combinator **A**. If AGR checks a [-Case] projection, then the head and its complement are combined by R-composition, and the semantic content of AGR is **R**. In 4.4, it is shown that the use of **R**-composition allows a lexical entry for np-trace to be given.

4.1 Raising as R-composition

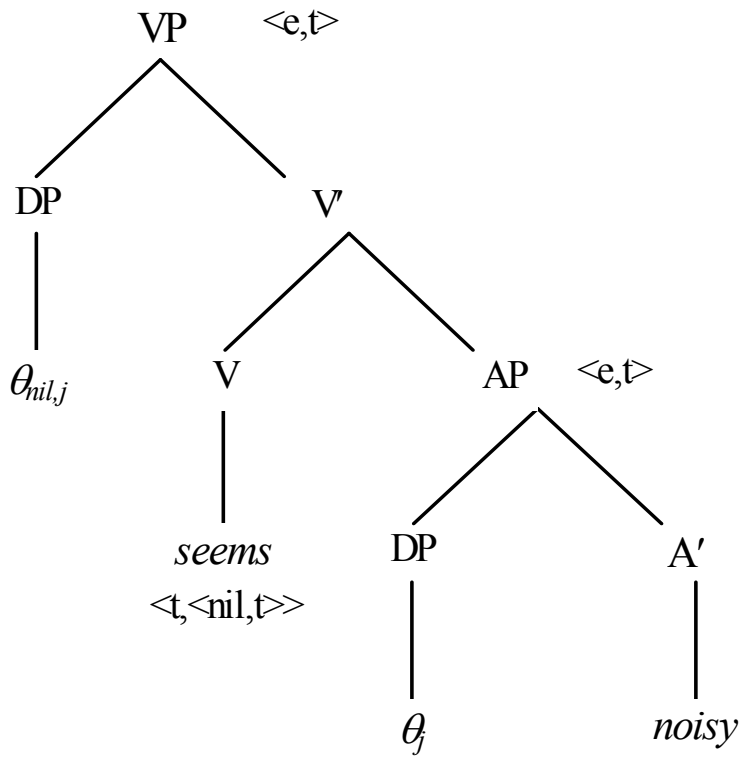
In this section, I will discuss the semantic composition of complex predicates. I have argued that a predicate is a constituent with an external θ-role unbound, rather than a constituent with an NP-trace in its specifier position. It follows from the distinction I have drawn between [+Case] and [-Case] determiners that a noun phrase as such cannot 'move' from a non-Case-licensed to a Case-licensed position: it is only possible for a [-Case] noun phrase to be generated in the former position, and a [+Case] noun phrase in the latter. In consequence, A-movement of noun phrases e.g. Raising, must be replaced by some compositional device which in a loose sense transmits θ-roles up the tree when there is no Case-licensing of an argument. I discuss Raising here, and np-trace in section 4.4.

I will show how this works by example, first giving a syntactic/geometric characterisation, and then providing the semantics to go with it. Let us take the raising verb *seems* as the head. I assume as standardly that *seems* s-selects for a proposition internally, but fails to Case license it; and I will suppose that the external argument is assigned a nil role, like *come* in (7). The type for *seems* will then be $<t,<\text{nil},t>>$. Consider then the sentence in (29):

(29) Every dog seems noisy

The XP predicate of the determiner *every* is the predicate IP [*seems noisy*]. It is clear that the actual θ-role bound by *every* comes from *noisy*. We can show this on the tree by spec to spec coindexing, as in (30):

(30)



I will show from the semantics that the coindexing is redundant: but it is nevertheless convenient as an indication of the way the θ-roles are transmitted up the tree to be discharged by a binder.

Now consider the c-selection and s-selection permitting this tree. The adjective I will assume to select simply for an external DP, of type <*e*>, so that its type is <*e,t*>, a predicate. The verb *seems* c-selects for AP internally, and for DP externally. Now *seems* cannot combine with the AP *noisy* by function-argument application: the AP is not an argument of the simple type <*e*> or <*t*>, nor is it a binder which can take the verb as its operand. It has the wrong type. But there must be some means by which the head and its complement combine, otherwise the constituent has no meaning. That is does have a meaning is intuitively obvious; and it can be the antecedent for VP anaphora, for instance, as in

(31) My terrier seems noisy, and so does Bill's Alsatian t_v [$_{VP}$ *e*]

I claim that [-Case] complements of lexical heads combine semantically with their selecting head by a slightly generalised form of function composition.[27]

If I had not postulated an external nil role for *seems*, this would have been the crossed function composition used by Jacobson (1990) in her *Raising as Function Composition.*[28] Essentially, function composition ascribes to the composed constituent as a whole the θ-role unsatisfied in the complement. As with the operand of *not*, discussed earlier, such a role will be interpreted as if *in situ*. Since the complement of *seems*, namely *noisy*, has an unsatisfied external argument, this role is ascribed to the composed constituent. However, under the nil-role analysis of ergatives, *seems* already has an external (nil) selection. I propose that the two roles are in a sense 'added' together: any argument discharging the first, discharges the second. The coindexing from the spec of the AP to the spec of the VP is intended to record just this fact.

The generalisation we want must ensure that the external roles from BOTH the head verb and its complement AP are ascribed to the whole. I will call this combination of the two functions, R-composition, since **R** is the related combinator. Algebraically, R-composition can be defined as follows:

(32) Let *f* be the function corresponding to the meaning of the head, and let *g* be the function corresponding to the meaning of the predicate selected. Then the R-composition of *f* with *g*, f_*g, is defined by

$$\mathbf{R}fg = f_*g = \lambda x\,([f(g(x))](x))$$

In this formula as applied to our example above, the first variable *x* marks the external argument of the complement AP, where a real role is given, and the second marks the external argument of *seems*, where the role is nil. Suppose (paralleling (8)) we take *seems ′* = $\lambda t \lambda y$ *Seems ′*(*t*), where *Seems ′* is the single-argument Language of Thought item, and the vacuous lambda-binding of *y* produces the nil role.. Then the semantics for [*seems noisy*] can be expanded as follows:

---

[27]Note that the composition here is obligatory, rather than being the optional composition used in categorial grammars to allow left to right processing and non-standard constituents. The latter is essentially a re-bracketing device, parasitic on the underlying expectation of function application. For the use of function composition in syntax and semantics, see Ades and Steedman (1982), Cormack (1989), Steedman (1992).

[28]The 'crossing' is a product of the categorial grammar assumption that subjects are selected leftwards and other arguments rightward.. Since we are not assuming that the compositional rules account for word order, we can just take this as composition, like the uncrossed version.

(33) **R** *seems´noisy´* = λx ([*seems'* (*noisy´*(*x*))](*x*))
= λx ([$\lambda t \lambda y$ *Seems´*(*t*) (*noisy´*(*x*))](*x*))
= λx (λy *Seems´*(*noisy´*(*x*)) (*x*))
= λx (*Seems´*(*noisy´*(*x*)))

The last line is obtained via vacuous quantification, and we have the expected meaning for the predicate.

We have now postulated that a lexical (relational) head and its complement can combine in one of two ways: by function argument application, or by R-composition. What licenses one or the other? The answer is clear: a Case-licensing head must combine by application; a non-Case-assigning head must combine by R-composition. If the types are not compatible with the given mode of combination, then the string is ill-formed. Case, then, enters into compositional semantics, as well as into lexical semantics.

4.2 Movement into a θ-position

I take it that the fact that the external θ-role assigned by the matrix head is nil cannot be accessed by the syntax: this is a matter of the semantic meaning postulates. If this is so, the use of R-composition implicitly violates the assertion that there is no movement into a theta-position (the MTC, Main Thematic Condition, in Brody's (1995) terminology; see also for example Chomsky (1981) and (1994) for discussion). The external argument in (29) bears two roles, one (nil) from *seems*, and the other by R-composition, mimicking movement, from *noisy*. I do not think that the MTC is correct. In my view it would be surprising if it were correct, since its function effectively is to limit the combinations of selections that a relational head can make. That is, a priori, we should suppose that a head could freely select externally for a nil role or a proper one, irrespective of the kind of complements and any Case licensing.[29] The version of the θ-criterion that precluded an argument from saturating more than one θ-role has been rejected since Chomsky (1986), removing the most obvious support for the MTC.

Brody (1995) constructs an argument involving parasitic gaps in support of the MTC, but it is not sufficient. The crucial examples are the following pair, where he argues that the (b) example is worse than the (a) example only because it violates the MTC:

[29]Burzio's generalisation must be rejected too (Burzio 1986).

(34) a) ?? Who did you believe *t* to have visited you without you having invited *e*

b) **Who did you believe *t* to have met everyone who invited *e*

It is not necessary to appeal to the MTC to explain the difference in acceptability. In (34b), the parasitic gap is in an adjunct to an NP (*-one*) within a DP, without there being a corresponding gap in the host NP. In (34a), properly, the gap is in an adjunct to the VP, where this host itself contains the primary gap.

Moreover, there is straightforward evidence that the MTC cannot be correct. Consider (35):

(35) a) The clothes [$\theta_{nil,nil,k}$ are [$_{AP}$ $\theta_{nil,k}$ easy [$_{CP}$ Op_k for you to iron t_k]]]
b) The clothes [$\theta_{nil,j,k}$ are [$_{AP}$ $\theta_{j,k}$ ready [$_{CP}$ Op_k for you to iron t_k]]]

In (35a), we see that the θ-role bound by [*the clothes*] is transmitted from the object position of *iron*. The CP is a predicate, whose external role Op_k derives from the object position of *iron*, as is intuitively clear. If we postulate that *easy* does not Case mark its complement, there will be composition, passing the role up to the spec of AP, which hold a nil role. Since *be* is not a Case-assigner either, the pair of roles are transmitted further, to be bound by the subject noun phrase. The problem arises with *ready* because the adjective, in (b), assigns two roles: internally, type <*t*>, and externally, type <*e*>. If we follow the pattern of (35a), the operator θ-role must be raised to the external position of the adjective - but this is already assigned a role, in violation of the MTC. The clausal complement of *ready*, which is unsaturated,[30] cannot be rendered saturated by the addition of PRO, as is done with Control structures, in order to side-step the problem

Similarly, there are problems with modals which have external arguments, since we expect these to select for VP. Possibly the only instance in English is deontic *must*, but Picallo (1990) shows clearly that in Catalan, *gosar*, 'to dare' is such a modal, and further argues that the modal does not select for a controlled clause. There will be external θ-roles as indicated in (36b), necessitating raising to a θ-position.

(36) a) En Joan li$_i$ gosava parlar t_i (Picallo's (33)b)
'Joan dared to talk to him/her'
b) En Joan li$_i$ [θ_1 gosava [θ_2 parlar t_i]]

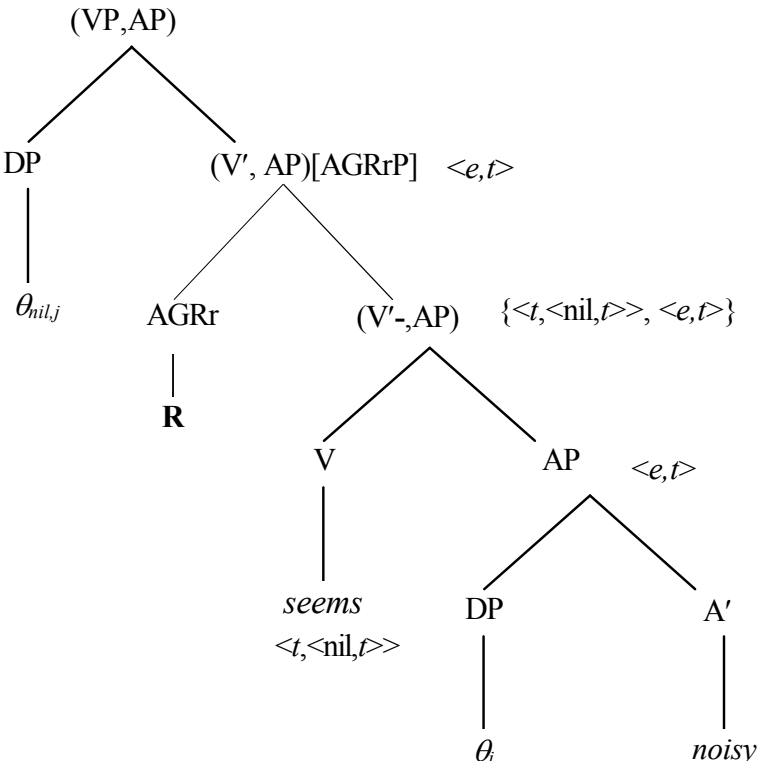
[30]The phrase arguably acts as a predicate in (i) *The fudge is for you to eat*

Higginbotham (1989) also argues that there are some cases of Raising in which the matrix verb already has an external θ-role. In Cormack (1995), I point to further MTC problems, relating to prepositional phrases, and I show that abandoning the MTC and using R-composition gives directly an alternative and very simple treatment of most cases of Control.

4.3 Linear ordering and AGR

We have not discussed linear ordering, so far. In section 3.2. I suggested for noun phrase objects that the noun phrase preceded the verb by virtue of the directional selection of the functional head D. But in (30), there is no functional head to determine the ordering. If we follow Chomsky (1981), the ordering should be indeterminate. Since in English, the verb must in fact precede the adjective, we must assume that the verb moves leftwards to some higher functional head. I suggest that this is a variety of AGR, since as we will see in a moment, it does check Case and φ-features, as AGR does in Chomsky (1993). We will call it AGRr, and assign to it the semantic content of the **R** combinator. In parallel, AGRo as in (27) will have the semantic content of the applicative combinator **A** (Szabolcsi 1990). The idea then will be that there can only be properly licensed θ-discharge if the appropriate combinator is present. Specifically, the double-headed projection which is the complement of AGR will be assumed to have a PAIR of functions, with their associated types, rather than a single function and type, for its denotation, as indicated in (37) below. This replaces (30), adding the intermediate AGRrP, so that the verb may move to AGRr.

(37)



AGRr will check that the verb *seems* is not a Case-assigner at the relevant projection (i.e. at its sister V′), so we will assume that there is a [-Case] feature present there, shown on the tree as '-'.

It might be supposed that the AGR head checks agreement between the verb *seems* and the adjective. Since the verb agrees with the subject, this might be expected to ensure that the adjective agrees with the subject, as is required in the French example (38), with two raising verbs.

(38) La fille semble être heureuse
the-FEM girl seems to be happy-FEM

However, we will find when we consider small clauses that this does not seem to be the locus of the agreement, but rather that the agreement is mediated by the higher AGR (AGRs, in this instance). In any case, such agreement induced by AGRr should be akin to object-agreement, so that if the verb at this point agrees with its complement, this would be nothing directly to do with the subject's φ-features.

4.4 np trace

Under the R-composition analysis above, there are no np-traces in subject position. A subject has to be a binder, and a binder has to be Case-licensed. Subject np-traces are impossible, as well as unnecessary. However, np-trace in a complement position is another matter. For one thing, there is c-selection to be satisfied; for another, the role assigned to this position has to be transmitted to some other position. Since R-composition does transmit roles, and np-trace is [-Case], we should hope to be able to identify the meaning and type of np-trace. It turns out that this is possible.

In the ordinary instances,[31] the required type for np-trace is <*e*,*e*>, since this will R-compose with a head which is expecting a type <*e*> argument. The meaning is just an identity function, λx (x).

(39) np-trace: category: DP selection: DP externally.
type: <*e*,*e*>; meaning: λ*x* (*x*)

To see how this works, consider the passive phrase [*seen t*]. Abstracting away from the problem of the proper representation of implicit arguments, we suppose *seen* to have a meaning encoding the fact that the usual external argument is existentially quantified, and the new external argument has nil semantic role:

(40) seen′ = λ*x*λ*y*[∃*p*, *p* see *x*], type <*e*,<nil,*t*>>

(41) seen′ * np-trace ≡ λ*x*λ*y*[∃*p*, *p* see *x*] * λ*x* (*x*)
≡ λ*z* ([λ*x*λ*y*[∃*p*, *p* see *x*] ([λ*x* (*x*)](*z*))](*z*))
≡ λ*z* ([λ*x*λ*y*[∃*p*, *p* see *x*] (*z*)](*z*))
≡ λ*z* ([λ*y*[∃*p*, *p* see *z*]](*z*))
≡ λ*z* ([[∃*p*, *p* see *z*]]) , type <*e,t*>

Thus R-composition of the nil-transitive verb with np-trace has produced a normal intransitive phrase, as required.

Only the initial, c-selected, np trace has a lexical entry. There are no intermediate traces, though it is possible as usual to show the transmission of the θ-roles under R-composition by coindexing the spec positions. For trace itself, no coindexing is required: the use of R-composition forces local transmission.

---

[31] A variant is required for the np trace following an unaccusative, when it is involved in covert conjunction with a transitive, since we want selection for a complement with a nil role. See Cormack and Breheny (1994) and Cormack and Smith (1994) for examples.

One of the reasons I cited in Cormack (1989) for abandoning the Internal Subject Hypothesis in favour of 'θ in spec' was the necessity for the interpretability of sub-sentential phrases such as passive VPs. The trace in these phrases will now be a legitimate entity, contributing to the compositional derivation of a predicate meaning. The introduction of a lexical entry for np-trace can be seen as part of a move toward the identification of D-structure, S-structure, and LF. The trace can be base-generated in situ, and because of the use of **R**, it can be properly interpreted, as is required of the elements of the interface LF. In a minimalist framework, np-trace will be a possible member of the initial enumeration. Similarly, the various members of AGR, which have semantic content, will be available in the initial enumeration.

5 Small Clauses

We are now in a position to give an account of the structure of small clauses which is consistent with the various claims we have made. Let us take

(42) Mary considers every dog noisy.

The verb *consider* s-selects for a proposition, which has as its head the adjective *noisy*. However, the noun phrase [*every dog*] is Case-licensed only by virtue of the Case-assigning property of *consider*. We assume the principle that every phrase must be locally licensed. Hence [*every dog noisy*], with [*every dog*] as a binder, cannot be a well-formed constituent. It is necessary, then, that *consider* combines first with *noisy* by R-composition, and that the noun phrase [*every dog*] be Case-licensed by this constituent.[32] The verb then moves to AGRo, as in (27) above, to give the surface order.

Suppose *consider* is simply of type $<t, <e,t>>$, with an internal propositional and an external entity argument. Then R-composition with *noisy* would require that the

[32]Such a constituent was proposed by Chomsky (1975, §101.1; but see §109.1), and di Sciullo and Williams (1986); Williams (1994) rejects this analysis. Chomsky (1986; 92) discusses the possibility, but rejects it on the grounds that if the main verb and the predicate jointly s-select the small clause subject, then this should not be able to be expletive. However, if s-selection includes selection for a nil role, there is no problem: R-composition will permit expletive subjects just when both the object role of the matrix verb, and the external role of the predicate, are nil, as in (43). Note that the structures effectively mimic those produced by reanalysis, so that arguments for the latter, such as those of Rizzi (1986, 535) can be taken as support for the proposals here. Vanden Wyngaerd (1994) has a gapping argument which requires that Small Clauses have a complex predicate analysis at some level.

external role of *noisy* be bound by the external argument of *consider*; which is not the correct result. It follows that *consider* must have a third role, an object role, which is semantically nil; it will be the binder of this which binds the external role of *noisy*. The type for *consider* is now <*t,*<nil,<*e,t*>>>, with c-selection for AP and DP internally. The c-selection is in line with both Williams (1994, §3.1.3) and Bowers (1993).

Essentially, R-composition will promote the semantic subject of the proposition into the empty object position, more or less as in the Raising to Object analyses of Emonds (1976), deriving from Rosenbaum (1967). It is the introduction of selection with nil semantic role that makes this analysis compatible with the principle that all structure is head-mediated, and the abandoning of the MTC which makes the subsequent 'movement' allowable.[33] Note too that as the verb now c- and s-selects for an object, then there is nothing surprising about the licensing of Case on this object, so that there is no reason to suppose that the Case is structural rather than inherent. The verb *consider* licenses Case when it projects its nil role, but does not license Case when it projects its type <*t*> role. It is only Case licensed by T for subjects which is structural, in the sense that (by assumption) T projects no θ-role, nil or otherwise, to be discharged by the subject.

The syntactic and semantic details will be spelled out in a moment, but first, the type offered needs to be shown to have some plausibility. Why is there an object with a nil role? Note that it cannot be any role other than the nil role that is assigned to this object position, because we may have sentences with expletive *it* in that position:

(43) Mary considers it absurd that the sky is blue

External arguments are assigned nil roles when the head deploys no θ-role associated with suitable meaning postulates - for example, no role with agentive properties. The relevant properties appear to vary from one language to another.[34] We may suppose that something similar dictates when the second position, that normally associated with the patient role, should be nil.

The tree for the VP [*considers every dog noisy*] will be as in (44), except that the movement of the verb to AGRr and then to AGRo has not been shown. The two intermediate levels of V-projection are distinguished as V′ and V″. The three θ-roles assigned by the verb are shown explicitly and superscripted. Of these, θ^2 is semantically nil. The A head is fully saturated and licensed where it is given as A*.

---

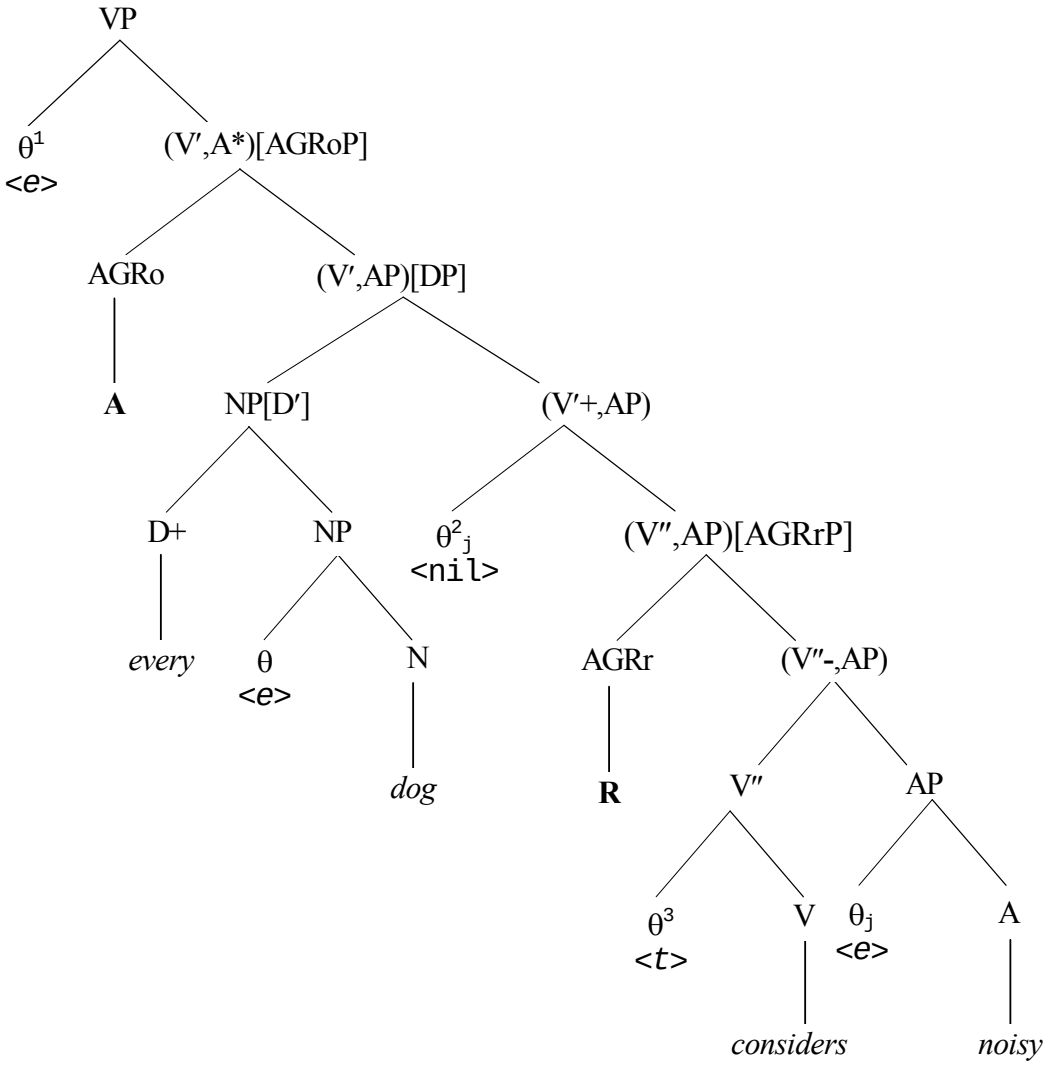
[33]Bowers (1993) and Authier (1991) also argue for Raising to Object, but on different grounds.

[34]See for instance van Hout et al (1993).

I have shown with a '+' the projection of the V which is a Case-licenser (V′), and also marked with a '+' the functional category D which must be Case-licensed in order to be a binder. AGRo checks for the [+Case] feature on D and for the [+Case] feature on the relevant projection of V; AGRr checks similarly that there is a [-Case] feature on the relevant projections of the selecting head (V) of its operand, and that if the other operand is headed by D, this too is [-Case].[35] It is clear that AGRrP must protect AGRo from looking down to the V" projection of V. Similar considerations would apply if the verb Case-marked two internal arguments, in order that the correct morphological case be associated with each argument. (The feature '+' in (44) is associated with [accusative]). This will be ensured if the relevant features are deleted as they are checked by AGR. If Case is associated with a particular projection of the head, rather than with the head as such, then we avoid any problem about the percolation of Case from inside the complex AGRrP to the binding noun phrase.

[35]But see section 6.3.

(44)



It will also be the case that AGRo checks agreement. It will check agreement between the various heads in its complements. These are AP and V′, and DP. What this means is that DP (or rather its functional head D) must subject-agree with A, and object-agree with V. [36]

The semantics is determined by the contents of the two AGR categories. **R** has the effect of discharging θ^3 with the content of the argument AP, while associating the undischarged θ_j role of the AP with θ^2. The combinator **A** allows NP[D′] to

[36]Note however that predicate A is ergative, so that there will be an object trace which object-agrees with A via another AGRr; what shows overtly on A could equally be subject or object agreement.

discharge the complex role θ^2_j. The semantics for **R** applied to a pair of constituents *f*, of type <*t*,<nil<*e*,*t*>>>, and *g*, of type <*e*,*t*> can be expanded as follows:[37]

(45) **R***fg* = $f_* g = \lambda x$ ([*f*(*g*(*x*))](*x*)) (=(32))
Let *f* = *considers′* = $\lambda P \lambda w \lambda y$ [(C*onsiders′ P*)(*y*)], where C*onsiders′* is the simple transitive Language of Thought item, and where *w* is for the nil role, which binds vacuously in the representation.
Let *g* = *noisy′* = λz[*noisy′*(*z*)]
R*fg* = λx ([*f*(*g*(*x*))](*x*))
= λx ([$\lambda P \lambda w \lambda y$ [(*considers′P*)(*y*)] (λz[*noisy′*(*z*)] (*x*))](*x*))
= λx ([$\lambda P \lambda w$ λy[(*considers′P*)(*y*)] (*noisy′*(*x*))](*x*))
= λx (λw λy[(*considers′*(*noisy′*(*x*)))(*y*)] (*x*))
= $\lambda x \lambda y$[(*considers′*(*noisy′*(*x*)))(*y*)]

The lambda-discharge to give the last line is vacuous. We have the required result, which is that the complex **R***fg* is just like a simple transitive projection, by means of exactly the **R** which we were forced to introduce as a variant of composition to deal with the Raising cases.[38] This result supports the claim that the combinator **R** is implicated in the semantics, and in the syntax in as much as theta theory belongs in the syntax. It also supports the claim that some form of Raising to Object is the correct characterisation of Small Clause constructions. ECM constructions will fall under the same process, with a VP[*to*] replacing the small clause predicate.[39]

The analysis offered here in a sense supports both the proponents and the opponents of Small Clauses. There is s-selection for a type <*t*> complement, just as there is for regular clauses, and the c-selected AP predicate as in the example above is duly saturated, by the argument which is taken to be the Small Clause subject. However, the saturation is not immediate, and the predicate and its external argument form a syntactic constituent only because of the movement of the verb.

The external role of the NP is discharged without the aid of any AGR: we may suppose that this is because the mode of discharge is not by binding (i.e. the type of the NP[D′] category is not <*t*>, as we would have if we bound the external role of the

[37]Combinators are properly unary operators. It is a simple matter to define binary operator equivalents.

[38]Jacobson (1990) too derives such complex predicates from composition, but the object-selection is derived entirely by composition, rather than partially from selection by the matrix verb as above.

[39]See Bowers (1993) for arguments to this effect within his framework.

<*e,t*> predicate NP, but the type for the generalised quantifier, <<*e,t*>,*t*>). Rather than cluttering up the tree with many more combinators, we may assume that every non-binding operator, like D as it applies to its operand NP, incorporates its own AGR combinator **A**.[40] This will enable us to account for possible overt agreement of D with NP.

Suppose now that the verb is passivised. Then NP[D′] will be replaced by np-trace, and there will be no [+] feature on V′. As we saw in section 3.2, the semantics of np-trace allows AGR to be instantiated as **R**, with the effect that the role θ^2_j is associated with θ^1, so that it will be bound by the subject, as is required. The c-selection by *consider* is unchanged under passive.

6 AGR heads

6.1 Summary

I have argued in the previous sections that Case - and in particular, the failure of a head to license [+Case] on a complement - necessitates a mode of composition of head and complement which is not the usual function-argument application. The analysis was motivated by considering Raising structures. Because of the introduction of nil roles in subject positions with no normal θ-role, the mode of composition cannot be ordinary function composition, as used by Jacobson (1990), but must be R-composition. R-composition effectively raises a θ-role to a position where there is already a θ-role assigned, adding the two roles together so that both are saturated by the same argument.

Note that the introduction of nil roles for external arguments leads to an alternative explanation for EPP effects. Every θ-role must be discharged if the result is to be a proposition. So every phrase which is to be construed as a proposition must have some subject to discharge the external role, even when this role is nil.

The linear order of head and predicate adjective in Raising constructions led to the postulation of another AGR, AGRr, to which the verb moves. AGRr takes on some of the functions of Chomsky's AGRa. It is suggested that AGRr is the locus of the R-composition combinator, **R**. Similarly, AGRo, (and as we shall see below, AGRs), has for its semantic content the applicative combinator, **A**. The choice of one or the other correlates directly with the Case-licensing properties of the relational head at the

[40] See Section 6.4 for further discussion.

associated projection, with **A** checking for [+Case] and **R** for [-Case]. Thus the instantiation of the compositional reflex of Case is in the content of AGR.

When Small Clause constructions were considered, it turned out that the compositional apparatus already set up gives a direct account, provided we accept a version of Raising to Object. It was argued that within the present framework, that is natural and harmless. ECM constructions are taken to parallel Small Clauses.

There is a sense in which the proposals do conform to Higginbotham's (1985) dictum that all arguments are saturated - a lexical head can only project θ-roles of primitive types, and these roles must be saturated.[41] However, the use of R-composition allows the saturation of such roles to be locally partial, which is not what Higginbotham envisaged.

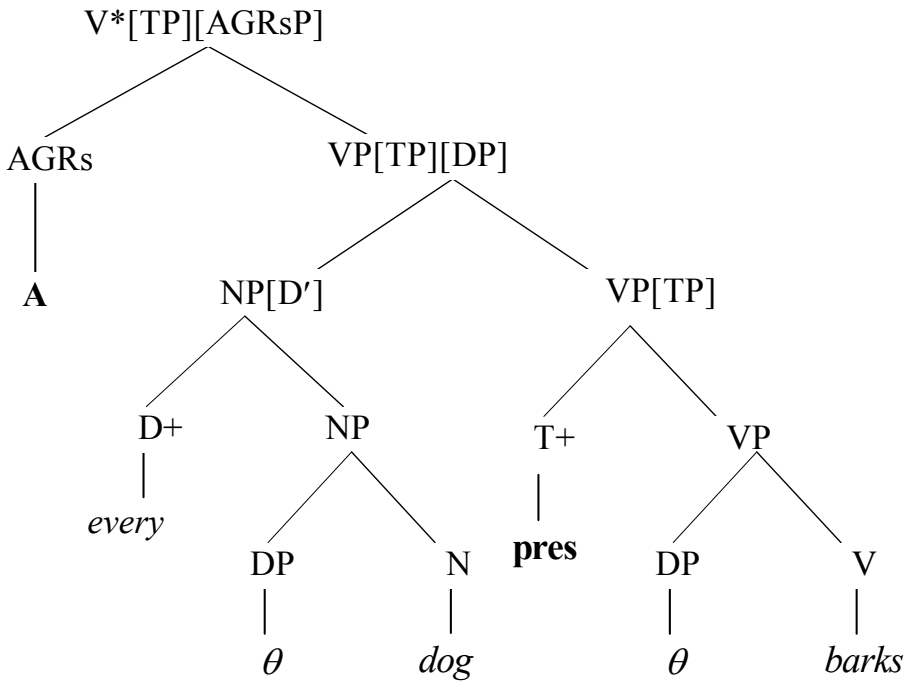
6.2 AGRs

Consistency requires that AGRs has as its complement the projection headed by the binding D of the subject, and the Case-licensing projection T.[42] For the sentence *every dog barks*, we must then have the tree in (46):

[41]Roles must be saturated only if the whole is to be a proposition - in free standing VPs, for instance, as we have mentioned before, the external role will not be saturated.

[42]Notice that the projection of the verb licensing the final (external) argument does not have a [+/–Case] feature. Predicates, then, are distinct from lower projections of a head in this respect.

(46)



There are two differences between AGRs and AGRo. First, the former checks for T[+Case], whereas the latter checks for a lexical projection with [+Case]. Second, it is necessary that the head move into AGRo in English, whereas it is not required that it move into AGRs.

VSO languages that are not V2 languages would have movement of lexical T, or of T+V, to AGRs.[43] This would replace the alternative more standard account of leaving S in spec of VP while V moves to T. Whether T+V moves to AGRs or not, the assumption is that [+Case] is checked by AGRs head-to head. In English, at least for some dialects, negative inversion might make use of AGRs:

(47) John said that never had he seen such a thing.

Alternatively, AGRs could be a property of C, or could be constrained to move to C. Under either of these suggestions, the existence of languages like West Flemish, with agreement between a subject and the complementiser of its clause, could be explained

[43]Picallo (1990) shows that at least in Catalan, epistemic modals are generated under T.

readily.[44] However, direct association of AGRs with C is not necessary for this. A functional head may have the option of agreeing with a head for which it selects: for example, in many languages, D must agree with the head of its first operand, NP. Then we might expect C to agree with T, with AGRs constraining T to agree with the subject D.

6.3 Case, m-case, and AGR

It may have occurred to the reader that there is a certain redundancy in the Case mechanisms I have been proposing. The question is whether once we have compositional content in AGR, it is necessary to have AGR checking D for +/- Case at all. Williams (1994) for instance claims that this can be left to take care of itself: a wrong choice will lead to ill-formedness. However, if we interpret Williams within a Case framework, we see that for arguments, he is relying on Burzio's generalisation - composition (i.e. Williams' relativised head mechanism) is required if and only if the external role of the head is empty. We cannot rely on that here because of the introduction of nil roles, whose nullity is inaccessible to syntax (and because of the generalisation to control structures). However, if the Case properties of the relational head at its various projections are part of the head's lexical entry, and these determine the content of the various AGR heads, then it is true that using the wrong entry for D would lead to a θ-theory crash: semantic interpretation would be unattainable. Let us suppose, then, that AGR does NOT check Case on D.[45] Phrases such as 'D[+Case]' are to be read as meaning, 'D suitable for combining via the combinator **A** licensed by a [+Case] head', and similarly for D[-Case] and **R**. The lexicon still contains more than one entry for many determiners, but without Case-features.

The system will now look somewhat more like that of Jacobson (1990), where the mode of combination is registered within the categorial representation of the head. The current interpretation removes an anomaly in the checking of complement heads: if the head is [+Case], there is always this feature to check, and also possibly there is a [-Case] feature on the D of np trace and Small Clause headed by D, but if the head selects say AP, there can be no [-Case] feature to check on the complement (see section 5).

---

[44]See Rustick (1991) for discussion of C-agreement.

[45]Under the proposals of the next section, the derivation would also crash because of failures of c-selection-feature checking and deletion (recollect from section 3.1 that argument and predicate Ds c-select differently).

It is still necessary that AGR check m-case agreement between complement and the projection of the head. For most languages, m-case is checked only if the selecting projection is [+Case]; this is the remaining link between [Case] and m-case. However, in Icelandic, for instance, it appears that m-case may be checked even when the selecting projection is [-Case]: this seems to be one of the defining characteristics of quirky case.

We may conclude that AGR checks that the m-case and φ-features on the heads of its complement match, subject to the various restrictions mentioned above; and that it also checks that the [Case] feature visible in its complement is consistent with its own combinator content.

6.4 AGR, merge, and the minimalist framework

The effect of AGRr is to reformulate some of the relationships usually characterised as np-movement as instances of a particular kind of combinatorial syntax and semantics. Suppose that all A-movement can be reformulated in the same way.[46] Then we should view the result within the minimalist framework as follows. The operation of Merge has two forms, depending on whether it is based on **A** or **R**. Most functional heads contain **A** as a feature at each non-maximal projection.[47] AGR may contain either **A** or **R**. However, the potential θ-discharging categories (D, C and P) are defective in this respect, so that merge must involve AGR. In addition, if no functional category is involved, as when a lexical head selects for another lexical head, merge must again involve AGR.

In order to account for head movement to AGR, we may postulate that each of the combinators in fact comes in two varieties, with a weak or strong L-attracting feature, so that movement of a proper lexical category (eg a head) to strong AGR is forced pre-PF.[48] It seems natural to characterise such movement as PF-movement. Minimally, one would not expect any LF movement: all the results of merge should be interpretable as they are constructed. In addition, the minimalist approach to composition suggests that all checking for selection requirements should be local, and

---

[46]Cormack (1995) discusses expletives, as well as saturation of θ-roles by CP and PP.

[47]This is to allow for two-place operators, which combine with each operand in turn by function-argument application i.e. with **A**.

[48]AGRa comes in two varieties in English, then: one is strong, and must check a lexical head; the other is a weak and must check T. AGRo is parametrically determined to be strong.

hence available at merge. The optimal configuration for checking dictates that it is head-to-head, with AGR checking accessible heads at merge.

In order to substantiate the idea that what has traditionally been seen as movement can be seen as Merge, I will compare standard Merge of a head and an argument, i.e. **A**-merge, with an instance of **R**-Merge, where in each instance the c-selection features are shown. I shall show these simply as features within square brackets, each preceded by a slash to distinguish them from features of other kinds. It is to be understood that when there is more than one, they are ordered in a stack, with the final category to be discharged first, in conformity with the usage of Categorial Grammars. Thus X[/Y/Z] is equivalent to (X[/Y])[/Z].

For the purposes of this discussion, I am assuming that a minimalist account of c-selection would disallow the selection of a complex category. So a head selecting for a CP will select for C; a head like D selecting for a predicate NP say will select simply for N, with the fact this must appear in the form of N[/D] ensured by the type system.[49] I will begin with a simplified story. H is a head, which selects as shown.

(48)	daughter 1	daughter 2	mother
a)	H[/A]	A	H
b)	H[/B/A]	A	H[/B]
c)	H[/A]	A[/D]	H[/D]
d)	H[/D/A]	A[/D]	H[/D]

In (48a), we have the simplest merge possible, registering that the c-selection has been discharged. Similarly, the discharge of the first complement of two would have the merge-tree in (b). These give simplified **A**-merge trees. Now consider (c). Suppose H is an ergative head. Then the external argument required at the mother is obtained from [/D] of the internal argument, A[/D], as in Raising - but what is raised is the feature, not an argument. This merge would be 'Raising as function composition'. However, I have been arguing in the preceding sections for a different form of composition, based on the notion that unaccusatives have nil external roles. There is necessarily selection for an external argument by H, then, as in (d). We

[49]The results then will differ from what would appear if a normal categorial grammar formalism had been used.

suppose that external selection is always DP. Then the D of H[/D/A] is an argument having a nil role.[50]

It seems to me that there is no principled difference between such forms as (a) and (b), and those like (c) or (d). The difference lies largely in the semantics. I contend, then, that it is proper to consider (c) or (d) too to be instances of Merge - and in particular (d) is simplified **R**-merge. In all cases, one selection feature from the head has been eliminated: an argument of the head has been discharged. The merge-trees of (a), (b) and (d) have in addition the property that the c-selection features of the mother include only those of the head. We could, if we wished, characterise Merge by these properties.

Under this interpretation, the operations of Williams' vertical binding, and Higginbotham's transmission of thematic roles (Higginbotham 1989) and the θ-transmission of Cormack (1989) are not to be seen literally as instance of 'move θ' - this notion, as Chomsky (1995) points out, makes little sense[51] - but as instances of Merge.[52]

The simplification alluded to above arises from the omission of the AGR nodes, and the ignoring of functional heads. In fact it will turn out that under the more detailed analysis, there is no difference at all in the c-merge operations required in a system with or without **R** in addition to **A**. There are five configurations to consider: one and two-place non-binding operators, two-place binding operators, lexical heads selecting for lexical projections, and AGR projections. Let us consider these in turn.

Suppose F is a one-place operator. Then the merge-tree will be simply as in (49):

---

[50]More properly, as I argue in Cormack (1995) in relation to control, the external argument roles of each of the daughters are combined by **R** and assigned to the mother, so that the role of D in H[/D/A] need not be nil.

[51]Movement of a theta role could be interpreted as movement of a feature, but this would in any case make incorrect predictions. Consider the VP anaphora in (i):

(i) The cat seems to like Mozart, but the baby doesn't [$_{VP}$ *e*].

There are two interpretations: the empty VP is construed to mean either 'seem to like Mozart' or 'like Mozart'. On the current analysis, these can both be obtained by copying VP[AGRP] phrases in the antecedent, or their meanings; the meanings will be of the type $<e,t>$ as required (the feature giving the tense morphology on *seems* must be dropped). Note that transmission of the θ-role is accomplished compositionally by the content of AGRr: the meanings of the internal phrases do not change. It is not the case that the external θ-role of *like* has actually moved and become unavailable in situ: this would make the inner VP meaningless, or lacking an external role.

[52]However, if we use spec positions on the tree to remind us of the content of the external θ-role, as I have suggested, then coindexing and a movement metaphor are generally helpful and harmless.

(49)	daughter 1	daughter 2	mother
	F[/A]	A	{A,F }

Here, we might for instance have F= *not*, and A = *cold*.[53] Note that the full description of the mother node includes the total categorial information about the daughters. The c-selection feature on the head can be inferred, since lexical categories cannot discharge their complements except under AGR. The category {A,F} is a lexical category, where A is lexical and F is a saturated one-place operator. The basic merge operation shown is to be understood as a pair-forming operation, not as set union.

Suppose F is a two-place operator, with operands A and B. The two levels of merge will be as in (50), (a) and (b):

(50)		daughter 1	daughter 2	mother
	a)	F[/B/A]	A	{A, F[/B]}
	b)	{A,F[/B]}	B	{B, {A,F}}

Here, we might have F = *because*, and A = B = IP.[54] At the second stage, we have a functional category, a one-place operator {A,F[/B]}. A category with an unsaturated functional head is functional; otherwise the category is lexical. The lexical category A of the pair is simply a feature of this functional category, and is inert with respect to the discharge of the feature [/B] at the next stage. Thus, this merge is obtained simply by two applications of the merge in (49).

The structures above represent what are more usually described as adjunction structures, for which Chomsky (1995) provides an arbitrary labelling, with the label consisting of a doubled head category. The discharge of the selection features is a variety of checking, under a simple sisterhood relation rather than a spec-head relation. The checked feature is deleted, allowing the next to become accessible.

Suppose now F is a binding operator - D. It will have a category D[/X/N]. Its first operand is necessarily N[/D], however, so if the first stage is as for a non-binding operator, we will have the merge-tree whose mother category is a derived one-place binding operator D′[/X], where D′[/X] = { D[/X], N[/D]}. The category D′[/X] is a generalised quantifier. Notice that under these assumptions, N is not yet saturated.

---

[53]Previously, we have had AP[FP] as the notation for the mother.

[54]The mother nodes would previously have been labelled F[IP_1] and F[IP_2].

We can interpret this as meaning that it is not until after both operands of the D head are available to D that a truth value can be computed.

The next stage(s) must be different, because θ-discharge is involved, and by hypothesis, this involves AGR. In particular, in the semantics, I postulated that there was initially simply a pair of meanings, with no θ-discharge. However, for the categories, I am going to assume that there is discharge of the selection by D for X. The merge is as shown in (51), where D′ represents the complex category {D,N} as above:

(51)	daughter 1	daughter 2	mother
	D′[/X]	X[/D]	{ D′, X[/D] }

This is just as in our other examples so far as categories are concerned.

At this point, AGR is called into play, and the further variety of merge, with the semantics given by **A**, takes place. The combinator operates on the pair given by its operand.[55] Notice that the argument D must now saturate both of the categories that select for it, when AGR acts as a catalyst for the discharge. The AGR merge will be as in (52):[56]

(52)	daughter 1	daughter 2	mother
	AGRa	{{ D, N[/D]]}, X[/D] }	{{X,{D,N}}, AGRa}

The category {{X,{D,N}, AGR} is a lexical category, and the major head is X, as required.

Let us consider next what happens when a lexical head selects for a lexical complement, say as in the selection for an AP small clause, *clever*, by *consider*. According to the discussion in section 5, the selections by the two lexical heads and the initial merge must be as in (53):

[55]The selection for AGR is [/{X,Y}] in every instance. I have omitted this for simplicity.

[56]I assume that we are dealing with an internal argument here; the extra T head will be needed for an external argument.

(53)	daughter 1	daughter 2	mother
	V[/D/D/A]	A[/D]	{ V[D/D/A], A[/D] }
	X	Y	{X, Y }

The result of applying AGRr to this, following the discussion in section 5, where I argued that the complement category must be preserved, is as in (54):

(54)		daughter 1	daughter 2	mother
	a)	AGRr	{ V[/D/D/A], A[/D]}	{<V[/D/D],A[/D]>, AGRr}
	b)	AGRr	{X[/Y], Y}	{{X, Y}, AGRr }

The semantics at this point is integrated, but the categorial description is not. We should ask whether the ordering in the pair is really necessary. It was imposed to conform to the intuition that the 'real' head of the phrase would be *consider*, not both *consider* and *clever*. We could reasonably argue that under the current interpretation of small clauses, the category assigning subject and object θ-roles is indeed the complex two-headed category *consider clever*. Let us replace the ordering brackets in (54a) by order-free ones, as in (54b).

When the category D object of *consider* is merged, using AGRa, we will get the following merge, modelled on that of (52):

(55)	daughter 1	AGRa
	daughter 2	{{ D, N[/D]]}, {{V[/D/D], A[/D]}, AGRr}}
	mother	{{{{V[/D], A}, AGRr},{D, N}}, AGRa}

Notice that the argument D has saturated the available argument-selection of all the lexical heads which were unsaturated. The simultaneous saturation of selections by the V and the A gives the control effect. At this point, the category A from *clever* is saturated, but the V from *consider* is not. This means that if we want to give primary heads only, dropping functional heads and complement heads as they are saturated, the A head from *clever* will be dropped, while the V head is retained. Thus the intuition that the primary lexical head is V has support.

Another instance of the discharge of two c-selections simultaneously is seen with type-shifted operators. Consider for instance a type-shifted $ (asymmetric conjunction) operator category, whose two operands are intransitive verbs. The category will be $\{V_1[D], \{V_2[D], \$\}\}$. When the argument D is supplied and AGRa-licensed, the result must be $\{\{V_1, \{V_2, \$\}\}, AGRa\}$. Thus in general, the effect of

AGRa is that the argument saturates the available c-selection from each non-saturated category of the (complex) head. Any saturated categories are ignored.[57]

The distinct merges postulated are those for an operator head, for a lexical head, and for an AGR head. These are shown again, schematically, below. That in (56), for an operator head, corresponds to what has been argued for above, in (49), (50), and (51). The merge-tree in (57a) for a lexical head corresponds to what was required in (53), and that in (57b) for an AGR head is as argued for in (52) and (54). The schema under (57b) is a simplification, since the argument Y may saturate not just the single [/Y] shown, but any number of other available selections of the feature within the complex lexical category, irrespective of the internal bracketing.

(56)	daughter 1	daughter 2	mother
	F[/Y]	Y	{F, Y }

(57)		daughter 1	daughter 2	mother
	a)	L	Y	{L, Y }
	b)	AGR	{L[/Y], Y}	{{L, Y}, AGR }

The trees as shown have the property that so far as the heads are concerned, the mother category contains the heads of the daughters. This means that Chomsky's notation giving a single head, and then the set of the daughters, as the full description of the category, is not necessary. All the information is available in the mother category. It is only if we turn to the abbreviated category labels, dropping saturated functional and complement heads, that information is lost.[58]

It perhaps should be reiterated that the exercise is purely one to assign category labels. The merges licensed in the type-system work in parallel, but not necessarily in one-one correspondence. In particular, the merge at the D′ level allows the discharge of selection of the one-place operator projection in the c-merge system, but allows no discharge in the s-selection system. However both systems use the pair-construction only when forced to: that is, at the level where there would otherwise be discharge of a lexical selection in the absence of AGR.

---

[57]Note that this is nothing to do with the types: the discharge of s-selection is strictly governed by the principles we have set up already, and it is always the case that one argument satisfies one θ-role of a unitary function.

[58]In at least the simpler instances, the selection information can also be retrieved, but whether this is always so is not clear without further work.

The conclusion which we can draw from the merges in (57) is that the postulation of a merge corresponding to **R** adds no extra complexity to the c-selection part of Merge. The form of merge required is, perhaps surprisingly, just the same for **R** as for **A**.

7 Conclusion

The Case-licensing feature of a projection determines two things. Primarily, it determines the semantic content of the AGR which checks this feature. Secondarily, the compositional syntax and semantics induced by AGR allows two forms of Merge, corresponding to the saturation of a θ-role, or to θ-transmission. If selection is for a DP, this in turn will determine the choice of lexical entry for the D involved. AGR, as well as having semantic content, may host head-movement. There are no A-chains.

References

Ades, Anthony, and Marc Steedman. 1982. On the order of words. *Linguistics and Philosophy* 4: 517-558.

Anderson, Stephen R. 1992. *A-morphous morphology*. Cambridge University Press,

Authier, J.-Marc. 1991. V-governed expletives, case theory, and the Projection Principle. *Linguistic Inquiry* 22: 721-740.

Barwise, Jon, and Robin Cooper. 1981. Generalised quantifiers and natural language. *Linguistics and Philosophy* 4: 159-219.

Borer, Hagit. 1984. *Parametric syntax*. Foris, Dordrecht.

Bowers, John. 1993. The syntax of predication. *Linguistic Inquiry* 24: 591-656.

Brody, Michael. 1995. Lexico-logical form: a radically minimalist theory. MIT Press, Cambridge, Mass.

Burzio, Luigi. 1986. *Italian Syntax*. Reidel, Dordrecht.

Chomsky, Noam. 1975. *Logical structure of linguistic theory*. Plenum, New York.

Chomsky, Noam. 1981. *Lectures on government and binding*. Foris, Dordrecht.

Chomsky, Noam. 1986. *Knowledge of language: Its nature, origin, and use*. Praeger, New York.

Chomsky, Noam. 1991. Some Notes on Economy of Derivation and Representation. In *Principles and parameters of comparative grammar*, ed. Robert Freidin, 417-454. MIT Press, Cambridge, Mass.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In *The view from building 20*, ed. K. Hale And S. J. Keyser, 1-52. MIT Press, Cambridge, Mass.

Chomsky, Noam. 1994. Bare phrase stucture. Ms., MIT.

Chomsky, Noam. 1995. Chapter 4. Ms., MIT.

Cormack, Annabel. 1989. The syntax and semantics of definitions. PhD dissertation, University of London.

Cormack, Annabel. 1995. The semantics of Case. Ms., in preparation, University College London.

Cormack, Annabel, and Richard Breheny. 1994. Projections for functional categories. In *UCL Working Papers in Linguistics* 6, 35-61. Department of Phonetics and Linguistics, University College London.

Cormack, Annabel, and Neil Smith. 1994. Serial Verbs. In *UCL Working Papers in Linguistics* 6, 62-88. Department of Phonetics and Linguistics, University College London.

Di Sciullo, A.M., and Edwin Williams. 1986. *On the definition of word*. MIT Press, Cambridge, Mass.

Emms, Martin. 1990. Polymorphic quantifiers. In *Edinburgh Working Papers in Cognitive Science* 5: 65-112.

Emonds, Joseph E. 1976. *A transformational approach to English syntax: root, structure-preserving and local transformations*. Academic Press, New York.

Giorgi, Alessandra. 1991. Prepositions, binding and theta marking. In *Long distance anaphora*, ed. Jan Koster and Eric Reuland ,185-208. Cambridge University Press.

Grimshaw, Jane. 1990. *Argument structure*. MIT Press, Cambridge, Mass.

Hellan, Lars. 1991. A two level X-bar system. *Lingua* 84: 239-257.

Higginbotham, James. 1985. On semantics. *Linguistic Inquiry* 16: 547-594.

Higginbotham, James. 1989. Elucidations of meaning. *Linguistics and Philosophy* 12: 465-517.

Ingham, Richard. 1991. Weak Islands, Trivial Pursuit and the ECP. Ms., University of Reading.

Jacobson, Pauline. 1990. Raising as function composition. *Linguistics and Philosophy* 13: 423-476.

Kayne, Richard S. 1994. *The antisymmetry of syntax*. MIT Press, Cambridge, Mass.

Ladusaw, William A. 1980. *Polarity sensitivity as inherent scope relations*. Garland.

Larson, Richard K. 1988a. On the Double Object construction. *Linguistic Inquiry* 19: 335-392.

Larson, Richard K. 1988b. Light Predicate Raising. Ms., MIT.

Manzini, M Rita. 1995. Bare dependencies. Paper presented at GLOW, University of Tromsø, Norway, June 1995.

May, Robert. 1985. *Logical form: its structure and derivation*. MIT Press, Cambridge, Mass.

Milsark, G L. 1977. Towards an explanation of certain peculiarities of existential constructions in English. *Linguistic Analysis* 3: 1-29.

Picallo, Carmen. 1990. Modal verbs in Catalan. *Natural Language and Linguistic Theory* 8: 285-312.

Rizzi, Luigi. 1986. Null objects in Italian and the theory of pro. *Linguistic Inquiry* 17: 501-557.

Rosenbaum, Peter S. 1967. *The grammar of English predicate complement constructions*. MIT Press, Cambridge, Mass.

Rustick, Susan. 1991. Incorporating a clausal head. In *Views on phrase structure*, ed. K. Leffel and D. Bouchard, 157-175. Kluwer, Dordrecht.

Saltarelli, Mario. 1988. *Basque.* Croom Helm, London, England.

Sperber, Dan, and Deirdre Wilson. 1986. *Relevance: Communication and cognition*. Blackwell, Oxford.

Steedman, Mark. 1988. Combinators and grammars. In *Categorial grammars and natural language structures*, ed. Richard T. Oehrle, Emmon Bach, and Deirdre Wheeler, 417-442. Reidel, Dordrecht.

Steedman, Mark. 1989. Constituency and coordination in a combinatory grammar. In *Alternative conceptions of phrase structure*, ed. Mark R. Baltin and Anthony S. Kroch, 201-231. Chicago University Press.

Steedman, Mark. 1992. Surface structure. In ICRS report no 92-32. University of Pennsylvania.

Stowell, Tim. 1981. Origins of phrase structure. PhD dissertation, MIT.

Stowell, Tim. 1991. Alignment of arguments in adjective phrases. In *Syntax and Semantics Volume 25: Perspectives on Phrase Structure*, ed. Susan Rothstein, 105-135. Academic Press, New York.

Szabolcsi, Anna. 1990. Combinatory Grammar and Projection From the Lexicon. In *Lexical Matters*, ed. Ivan Sag and Anna Szabolcsi, 241-268. CSLI Publications, Stanford University, Ca.

Van Hout, Angeliek, Janet Randall, and Jürgen Weissenborn. 1993. Are linking rules universal?: Learning unaccusativity in Dutch and German. Paper presented at Generative Approaches to Language Acquisition, University of Durham, England, September 1993.

Vanden Wyngaerd, Guido. 1993. Gapping, verb raising, and Small Clauses. Ms., K.U. Brussels.

Williams, Edwin. 1994. *Thematic structure in syntax*. MIT Press, Cambridge, Mass.